

*Proceedings of the 27th International Conference on Probabilistic, Combinatorial
and Asymptotic Methods for the Analysis of Algorithms*
Kraków, Poland, 4-8 July 2016

Bootstrap percolation on $G(n, p)$ revisited

Mihyun Kang[†], Tamás Makai[‡]

Graz University of Technology, Institute of Discrete Mathematics, Steyrergasse 30, 8010 Graz, Austria
{kang, makai}@math.tugraz.at

Bootstrap percolation on a graph with infection threshold $r \in \mathbb{N}$ is an infection process, which starts from a set of initially infected vertices and in each step every vertex with at least r infected neighbours becomes infected. We consider bootstrap percolation on the binomial random graph $G(n, p)$, which was investigated among others by Janson, Łuczak, Turova and Valier (2012). We improve their results by strengthening the probability bounds for the number of infected vertices at the end of the process.

Keywords: Random graph, Bootstrap percolation, Martingale

1 Introduction

Bootstrap percolation on a graph with infection threshold $r \in \mathbb{N}$ is a deterministic infection process which evolves in rounds. In each round every vertex has exactly one of two possible states: it is either infected or uninfected. We denote the set of initially infected vertices by $A(0)$. In each round of the process every uninfected vertex v becomes infected if it has at least r infected neighbours, otherwise it remains uninfected. Once a vertex has become infected, it remains infected forever. The final infected set is denoted by A_f .

Bootstrap percolation was introduced by Chalupa, Leath, and Reich [CLR79] in the context of magnetic disordered systems. Since then bootstrap percolation processes (and extensions) have been used to describe several complex phenomena: from neuronal activity [Ami10, ELP⁺] to the dynamics of the Ising model at zero temperature [FSS02].

In the context of social networks, bootstrap percolation provides a prototype model for the spread of ideas. In this setting infected vertices represent individuals who have already adopted a new belief and a person adopts a new belief if at least r of his acquaintances have already adopted it.

On the d -dimensional grid $[n]^d$ bootstrap percolation has been studied by Balogh, Bollobás, Duminil-Copin, and Morris [BBDCM12], when the initial infected set contains every vertex independently with probability p . For the size of the final infection set they showed the existence of a sharp threshold. More precisely, they established the threshold probability p_c , such that if $p \leq (1 - \varepsilon)p_c$, then the probability that every vertex in $[n]^d$ becomes infected tends to 0, as $n \rightarrow \infty$, while if $p \geq (1 + \varepsilon)p_c$, then the probability that every vertex in $[n]^d$ becomes infected tends to one, as $n \rightarrow \infty$.

[†]Supported by Austrian Science Fund (FWF): P26826, W1230.

[‡]Supported by Austrian Science Fund (FWF): P26826.

Bootstrap percolation has also been studied for several random graph models. For instance Amini and Fountoulakis [AF14] considered the Chung-Lu model [CL02] where the vertex weights follow a power law degree distribution and the presence of an edge $\{u, v\}$ is proportional to the product of the weights of u and v . Taking into account that in this model a linear fraction of the vertices have degree less than r and thus at most a linear fraction of the vertices can become infected, the authors proved the size of the final infected set A_f exhibits a phase transition.

Janson, Łuczak, Turova, and Vallier [JŁTV12] analysed bootstrap percolation on the binomial random graph $G(n, p)$, a graph with vertex set $[n] := \{1, 2, \dots, n\}$ where every edge appears independently with probability $p = p(n)$, and the set of initially infected vertices $A(0)$ is chosen uniformly at random from the vertex sets of size a . For $r \geq 2$ and p satisfying both $p = \omega(n^{-1})$ and $p = o(n^{-1/r})$, they showed, among other results, that with probability tending to one as $n \rightarrow \infty$ either only a few additional vertices are infected or almost every vertex becomes infected. In addition they determined, depending on the number of initially infected vertices, the probability of both of these events up to an additive term tending to zero as $n \rightarrow \infty$.

The main contributions of this paper are threefold. First we strengthen this result by showing exponential tail bounds. Second we introduce a martingale in order to determine the number of infected vertices during the early stages of the process. Finally in the supercritical regime we show that the subgraph spanned by the vertices with $r - 1$ infected neighbours grows large enough to contain a giant component. The infection of just one vertex in this giant component leads to every vertex in the component becoming infected and we show that this in fact happens.

Main Results. Throughout the paper we assume that $r \geq 2$ and that both $p = \omega(n^{-1})$ and $p = o(n^{-1/r})$ hold. Set

$$t_0 := \left(\frac{r!}{np^r} \right)^{1/(r-1)}.$$

Let $\hat{\pi}(t) = \mathbb{P}[\text{Bin}(t, p) \geq r]$ and define

$$a_c := -\min_{t \leq t_0} \frac{n\hat{\pi}(t) - t}{1 - \hat{\pi}(t)}.$$

In addition denote by t_c the smallest value t where this minimum is reached. Similarly to [JŁTV12] it can be shown that

$$t_c = (1 + o(1))((r-1)!/(np^r))^{1/(r-1)} \quad \text{and} \quad a_c = (1 + o(1))(1 - 1/r)t_c.$$

Theorem 1 *Let ω_0 be any function satisfying the conditions $\omega_0 = \omega(\sqrt{a_c})$ and $\omega_0 \leq a_c - r$. If $|A(0)| = a_c - \omega_0$, then with probability at least*

$$1 - \exp\left(-\frac{\omega_0^2}{10t_0}\right)$$

we have $|A_f| < t_c$.

Theorem 2 *Let ω_0 be any function satisfying the conditions $\omega_0 = \omega(\sqrt{a_c})$ and $\omega_0 \leq t_0 - a_c$. If $|A(0)| = a_c + \omega_0$, then with probability at least*

$$1 - \exp\left(-\frac{\omega_0^2}{10t_0}\right) - \exp\left(-\frac{a_c + \omega_0}{4}\right)$$

we have $|A_f| = (1 + o(1))n$.

Proof Technique. When the number of infected vertices is small (at most t_0), we introduce a martingale to show that the number of infected vertices is concentrated around its expectation with *exponentially* high probability. The martingale resembles the one introduced in [JLTV12], however the maximal one step difference in our martingale is significantly smaller and thus provides a tighter concentration bound (Lemma 7).

In the subcritical regime, the expected number of infected vertices is less than $t_c < t_0$ and therefore the martingale argument alone implies the result (Section 4).

In the supercritical regime, this is not enough as the number of infected vertices will reach t_0 with exponentially high probability. In fact, at least $t_0 + a_c$ vertices become infected (Lemma 8). Now take a subset of the infected vertices with size t_0 and consider the vertices with at least $r - 1$ neighbours in this set. The size of this set is roughly rp^{-1} (Lemma 9) and the subgraph spanned by these vertices is also a binomial random graph, $G(rp^{-1}, p)$. Since the seminal work of Erdős and Rényi [ER60], it is known that this graph has with probability $1 + o(1)$ a linear sized giant component. More recently, Bollobás and Riordan [BR] showed that this happens with exponentially high probability (Theorem 5). Should any vertex in the giant component have an additional infected neighbour, then every vertex in the giant will become infected eventually. We show that this happens with exponentially high probability.

Thus we have $\Omega(p^{-1})$ infected vertices. After this, the process ends in two steps and this can be shown by two simple applications of the Chernoff bound (Lemmas 10 and 11).

2 Preliminaries

We will use the following form of the Chernoff bound.

Theorem 3 [CL06] *Let $X \sim \text{Bin}(n, p)$, i.e. a binomial random variable with parameters n and p . Then for any $\lambda > 0$*

$$\mathbb{P}[X - \mathbb{E}(X) \leq -\lambda] \leq \exp\left(-\frac{\lambda^2}{2\mathbb{E}(X)}\right) \quad \text{and} \quad \mathbb{P}[X - \mathbb{E}(X) \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{2(\mathbb{E}(X) + \lambda/3)}\right).$$

Let M_0, M_1, \dots, M_i be a sequence of random variables and denote by $\mathcal{F}(i)$ the filter generated by M_0, \dots, M_i . We say M_0, \dots, M_k forms a martingale if for every $0 \leq i \leq k$ we have $\mathbb{E}(|M_i|) < \infty$ and for every $1 \leq i \leq k$

$$\mathbb{E}[M_i | \mathcal{F}(i-1)] = M_{i-1}.$$

The following concentration bound on martingales due to Chung and Lu [CL06] will prove to be vital.

Theorem 4 [CL06] *For $m_0 \in \mathbb{R}$ let $M_0 = m_0, M_1, \dots, M_k$ be a martingale whose conditional variance and differences satisfy the following: for each $1 \leq i \leq k$,*

- $\text{Var}[M_i | M_{i-1}, \dots, M_0] \leq \sigma_i^2$;
- $|M_i - M_{i-1}| \leq m$ for some positive m .

Then for any $\lambda > 0$, we have

$$\mathbb{P}[M_k - M_0 \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{2\left(\sum_{i=1}^k \sigma_i^2 + m\lambda/3\right)}\right).$$

We will also need the following Theorem on the appearance of a giant component in $G(n, p)$ by Bollobás and Riordan [BR].

Theorem 5 [BR] *Let $c > 1$ be a constant independent of n and let $\varepsilon > 0$ independent of n . Then with probability $1 - \exp(-\Omega(n))$ the binomial random graph $G(n, c/n)$ has a component of size at least $(1 - \varepsilon)pn$, where $\rho \in (0, 1)$ is the unique positive solution of $1 - \rho = \exp(-c\rho)$.*

3 Setup: Martingale

In order to analyse the bootstrap percolation on $G(n, p)$ we will use the following reformulation due to Scalia-Tomba [ST85] as in [JLT12]. Roughly speaking they examine the infected vertices one by one and determine the vertices which have at least r neighbours in the set of previously examined vertices. The set of examined vertices until step t is denoted by $Z(t)$ and the set of infected vertices by $A(t)$. Formally let $A(0)$ be the set of initially infected vertices of size a and without the loss of generality we may assume that $A(0) = \{1, \dots, a\}$. Set $Z(0) = \emptyset$. For each step $t \in \mathbb{N}$, if $A(t-1) \setminus Z(t-1) \neq \emptyset$, then let $U_t = \{u_t\}$, where u_t is a vertex in $A(t-1) \setminus Z(t-1)$ selected according to an arbitrary rule, otherwise set $U_t = \emptyset$. Set $Z(t) := Z(t-1) \cup U_t$. Now for $t \geq 0$ and each $i \in [n-a] := \{1, \dots, n-a\}$ let $X(t, i)$ be the indicator random variable for the event that the vertex $a+i$ has at least r neighbours in $Z(t)$ and set

$$A(t) := A(0) \cup \{a+i : X(t, i) = 1, i \in [n-a]\}.$$

The process stops when $t = n$.

Clearly $Z(t) \subset A(t)$. Let T denote the smallest value of t such that $A(t) = Z(t)$. Note that $t \leq T$ implies that $|Z(t)| = t$ and thus T is also the smallest t such that $|A(t)| = t$. Since $|A(t)| \leq n$ for every natural number $0 \leq t \leq n$ we have that $T \leq n$. Note further that $A(T) = A_f$.

In order to have a better control on the maximal number of vertices which can become infected in a single step, we refine the process by dividing every step into rounds, in such a way that in each round *exactly one vertex* $v \in [n] \setminus A(0)$ is examined (regardless whether it was examined in earlier rounds or not). Thus each step $1 \leq t \leq n$ consists of $n-a$ rounds and round i of step t is denoted by (t, i) . We denote the step following (t, i) by $(t, i) + 1$ and the preceding step by $(t, i) - 1$. Also the ordering of the rounds is given by the lexicographical order i.e. $(\tau, \iota) < (t, i)$ if either $\tau < t$ or $\tau = t$ and $\iota < i$.

In round i of step t we examine if vertex $a+i$ has at least r neighbours in $Z(t)$ and if it has we add it to the set of infected vertices. Formally for $(t, i) \geq (1, 1)$

$$A((t, i) + 1) := A(0) \cup \{a+j : j \leq i, X(t, j) = 1\} \cup \{a+j : j > i, X(t-1, j) = 1\}.$$

Clearly we have $A(t) = A(t, n-a)$. For consistency define $A(0, n-a) := A(0)$.

Define a function $\pi : \mathbb{N} \rightarrow [0, 1]$ by

$$\pi(t) := \begin{cases} \mathbb{P}[\text{Bin}(t, p) \geq r], & \text{for } t \leq T \\ \mathbb{P}[\text{Bin}(T, p) \geq r], & \text{for } t > T \end{cases}$$

and note that $\pi(t)$ is a random variable.

For $(t, i) \geq (0, n-a)$, define the random variable

$$M(t, i) := \sum_{j=1}^i \frac{X(t, j) - \pi(t)}{1 - \pi(t)} + \sum_{j=i+1}^{n-a} \frac{X(t-1, j) - \pi(t-1)}{1 - \pi(t-1)}. \quad (1)$$

We will denote by $\mathcal{F}(t, i)$ the filter generated by $M(0, n-a), \dots, M(t, i)$.

Lemma 6 *The sequence of random variables $M(0, n-a), \dots, M(n, n-a)$ forms a martingale.*

Proof: Fix $1 \leq t \leq n$ and $1 \leq i \leq n-a$. For every $\tau < t$, we can express from (1) the number of infected vertices in step τ :

$$|A(\tau)| = a + \sum_{\iota=1}^{n-a} X(\tau, \iota) \stackrel{(1)}{=} a + M(\tau, n-a)(1 - \pi(\tau)) + (n-a)\pi(\tau). \quad (2)$$

Recall that $\hat{\pi}(t) = \mathbb{P}[\text{Bin}(t, p) \geq r]$. We will denote by T' the smallest value of t which satisfies $a + M(t, n-a)(1 - \hat{\pi}(t)) + (n-a)\hat{\pi}(t) = t$. Because $\pi(t) = \hat{\pi}(t)$ when $t \leq T$, we have $T = T'$. Given the filter $\mathcal{F}((t, i) - 1)$ one can establish if $a + M(\tau, n-a)(1 - \hat{\pi}(\tau)) + (n-a)\hat{\pi}(\tau) = \tau$ for some $\tau < t$. Therefore, it can be determined whether the event $T' < t$ or $t \geq T'$ holds. In particular, if $T' < t$, then the exact value of T' can be determined.

For each $\tau \leq t$, since $\pi(\tau)$ depends only on the value of $T = T'$, we can also determine the value of $\pi(\tau)$, i.e. $\mathbb{E}[\pi(\tau) | \mathcal{F}((t, i) - 1)] = \pi(\tau)$ for $\tau \leq t$.

Note that $X(0, i) = 0$ for every $1 \leq i \leq n-a$ and that for every $(1, 1) \leq (\tau, \iota) < (t, i)$ we can easily compute from (1)

$$M(\tau, \iota) - M((\tau, \iota) - 1) = \frac{X(\tau, \iota) - \pi(\tau)}{1 - \pi(\tau)} - \frac{X(\tau - 1, \iota) - \pi(\tau - 1)}{1 - \pi(\tau - 1)}. \quad (3)$$

Therefore, based on the filter $\mathcal{F}((t, i) - 1)$, the value of $X(\tau, \iota)$ can be determined for every $(\tau, \iota) < (t, i)$.

Next we shall show that

$$\mathbb{E} \left[\frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1) \right] = \frac{\mathbb{E}[X(t, i) | \mathcal{F}((t, i) - 1)] - \pi(t)}{1 - \pi(t)} = \frac{X(t-1, i) - \pi(t-1)}{1 - \pi(t-1)}. \quad (4)$$

To this end, observe that if $X(t-1, i) = 1$, then we have $X(t, i) = 1$ with probability 1 and in this case both sides of equation (4) equal 1.

Now assume that $X(t-1, i) = 0$. When $t > T$, we have $X(t, i) = X(t, i-1) = 0$ with probability 1 and by the definition of $\pi(t)$ we have $\pi(t) = \pi(t-1) = \hat{\pi}(T)$. Evaluating both sides of equation (4) gives us $-\hat{\pi}(T)/(1 - \hat{\pi}(T))$. When $t \leq T$ and $X((t, i) - 1) = 0$, we have $\pi(t) = \hat{\pi}(t)$ and thus

$$\mathbb{P}[X(t, i) = 0 | \mathcal{F}((t, i) - 1)] = \frac{1 - \hat{\pi}(t)}{1 - \hat{\pi}(t-1)} = 1 - \frac{\hat{\pi}(t) - \hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)}.$$

Therefore in this case

$$\mathbb{E} \left[\frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1) \right] = -\frac{\hat{\pi}(t)}{1 - \hat{\pi}(t)} \frac{1 - \hat{\pi}(t)}{1 - \hat{\pi}(t-1)} + 1 \cdot \frac{\hat{\pi}(t) - \hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)} = -\frac{\hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)}$$

and thus (4) holds. According to (3) we have that

$$\begin{aligned} & \mathbb{E}[M(t, i) - M((t, i) - 1) | \mathcal{F}((t, i) - 1)] \\ &= \mathbb{E} \left[\frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1) \right] - \frac{X(t-1, i) - \pi(t-1)}{1 - \pi(t-1)} \stackrel{(4)}{=} 0. \end{aligned}$$

□

Lemma 7 Let $t \in \{0, \dots, n\}$ and $\lambda \in \mathbb{R}^+$ be given.

$$\mathbb{P} \left[\bigwedge_{(0, n-a) \leq (\tau, i) \leq (t, n-a)} M(\tau, i) > -\lambda \right] \geq 1 - \exp \left(-\frac{\lambda^2(1 - \hat{\pi}(t))^3}{2(n\hat{\pi}(t) + \lambda/3)} \right)$$

and

$$\mathbb{P} \left[\bigwedge_{(0, n-a) \leq (\tau, i) \leq (t, n-a)} M(\tau, i) < \lambda \right] \geq 1 - \exp \left(-\frac{\lambda^2(1 - \hat{\pi}(t))^3}{2(n\hat{\pi}(t) + \lambda/3)} \right).$$

Proof: We will only show the bound on the probability that $M(\tau, i) < \lambda$ for each $(\tau, i) \leq (t, n-a)$. The other case follows simply from the fact that if the random variables $M(0, n-a), \dots, M(t, n-a)$ form a martingale, then $-M(0, n-a), \dots, -M(t, n-a)$ is also a martingale and they both have the same conditional variance and maximal difference. In order to show that the bounds hold for each round, we introduce the following martingale:

$$\hat{M}(\tau, i) = \begin{cases} M(\tau, i) & \text{if } \hat{M}((\tau, i) - 1) < \lambda \\ \hat{M}((\tau, i) - 1) & \text{otherwise.} \end{cases}$$

Similarly to $M(t, i)$ we denote the filter generated by $\hat{M}(0, n-a), \dots, \hat{M}(t, i)$ with $\hat{\mathcal{F}}(t, i)$. Note that if there exists a round (τ, i) such that $M(\tau, i) \geq \lambda$, then we have $\hat{M}(\tau', i') \geq \lambda$ for every $(\tau', i') \geq (\tau, i)$. Therefore $\hat{M}(t, n-a) < \lambda$ implies that for every $(\tau, i) \leq (t, n-a)$ we have $M(\tau, i) < \lambda$.

By (3) and since $\pi(\tau) \leq \hat{\pi}(\tau)$ with probability 1, we have

$$|\hat{M}(\tau, i) - \hat{M}((\tau, i) - 1)| \leq \max \left\{ \frac{\hat{\pi}(\tau)}{1 - \hat{\pi}(\tau)} - \frac{\hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau - 1)}, 1 + \frac{\hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau - 1)} \right\} = \frac{1}{1 - \hat{\pi}(\tau - 1)}.$$

Note that $M(0, n-a) = 0$. Since $\text{Var}[\hat{M}(\tau, i) | \hat{\mathcal{F}}((\tau, i) - 1)] = 0$ if $\hat{M}((\tau, i) - 1) \geq \lambda$ and

$$\text{Var}[\hat{M}(\tau, i) | \hat{\mathcal{F}}((\tau, i) - 1)] = \text{Var}[M(\tau, i) | \mathcal{F}((\tau, i) - 1)]$$

otherwise, Theorem 4 implies that

$$\mathbb{P}[\hat{M}(t, n-a) \geq \lambda] \leq \exp \left(-\frac{\lambda^2}{2(S + \lambda/(3(1 - \hat{\pi}(\tau - 1))))} \right)$$

where

$$S \leq \sum_{\tau=1}^t \sum_{i=1}^{n-a} \text{Var}[M(\tau, i) | \mathcal{F}((\tau, i) - 1)]. \quad (5)$$

Note that

$$\begin{aligned} \text{Var}[M(\tau, i) | \mathcal{F}((\tau, i) - 1)] &= \text{Var} \left[\frac{X(\tau, i)}{1 - \pi(t)} \middle| \mathcal{F}((\tau, i) - 1) \right] \\ &\leq \frac{1}{(1 - \hat{\pi}(t))^2} \text{Var}[X(\tau, i) | \mathcal{F}((\tau, i) - 1)]. \end{aligned} \quad (6)$$

Recall that $X(\tau - 1, i) = 1$ implies $X(\tau, i) = 1$ and that $\tau > T$ implies $X(\tau, i) = X(\tau - 1, i)$. In both of these cases we have

$$\text{Var}[X(\tau, i) | \mathcal{F}((\tau, i) - 1)] = 0 \stackrel{\hat{\pi}(t) \geq \hat{\pi}(t-1)}{\leq} \frac{\hat{\pi}(t) - \hat{\pi}(t-1)}{1 - \hat{\pi}(t)}. \quad (7)$$

Now assume $\tau \leq T$ and $X(\tau - 1, i) = 0$. Since $X(t, i)$ is an indicator random variable, we have

$$\text{Var}[X(\tau, i) | \mathcal{F}((\tau, i) - 1)] \leq \mathbb{E}[X(\tau, i) | \mathcal{F}((\tau, i) - 1)] = \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau - 1)} \leq \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau)}. \quad (8)$$

Putting (5)-(8) together, we obtain

$$S \leq \sum_{\tau=1}^t \sum_{i=1}^{n-a} \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{(1 - \hat{\pi}(\tau))^3} \leq \sum_{\tau=1}^t \frac{n(\hat{\pi}(\tau) - \hat{\pi}(\tau - 1))}{(1 - \hat{\pi}(\tau))^3} \leq \frac{n\hat{\pi}(t)}{(1 - \hat{\pi}(t))^3}.$$

□

The previous lemma allows us to analyse the process in the first t_0 steps. This will be used in the proofs of Theorems 1 and 2.

4 Proof of Theorem 1

We want to investigate the number of infected vertices at time t_c . By the definition of a_c and t_c , we have

$$a_c = -\min_{t \leq t_0} \frac{n\hat{\pi}(t) - t}{1 - \hat{\pi}(t)} = \frac{t_c - n\hat{\pi}(t_c)}{1 - \hat{\pi}(t_c)}. \quad (9)$$

By the definition of $M(t, i)$, we have

$$|A(t_c)| \stackrel{(2)}{=} a + (1 - \pi(t_c))M(t_c, n - a) + (n - a)\pi(t_c).$$

Since $\pi(t) \leq \hat{\pi}(t)$ and $a = a_c - \omega_0$, we obtain

$$\begin{aligned} |A(t_c)| &\leq a + M(t_c, n - a) + (n - a)\hat{\pi}(t_c) \\ &= (a_c - \omega_0)(1 - \hat{\pi}(t_c)) + n\hat{\pi}(t_c) + M(t_c, n - a) \\ &\stackrel{(9)}{=} t_c - n\hat{\pi}(t_c) + n\hat{\pi}(t_c) - \omega_0(1 - \hat{\pi}(t_c)) + M(t_c, n - a) \\ &\stackrel{\hat{\pi}(t_c) \leq \hat{\pi}(t_0)}{\leq} t_c - \omega_0(1 - \hat{\pi}(t_0)) + M(t_c, n - a). \end{aligned} \quad (10)$$

Using $np = \omega(1)$ and $t_0 = (r!/(np^r))^{1/(r-1)}$, we have

$$t_0 p = O\left(\left(\frac{1}{np^r}\right)^{1/(r-1)} p\right) = O\left(\left(\frac{1}{np}\right)^{1/(r-1)}\right) = o(1). \quad (11)$$

Furthermore,

$$\begin{aligned}\hat{\pi}(t_0) &= \mathbb{P}[\text{Bin}(t_0, p) \geq r] = \sum_{j=r}^{t_0} \binom{t_0}{j} p^j (1-p)^{t_0-j} \stackrel{(11)}{=} (1+o(1)) \frac{t_0^r p^r}{r!} \\ &= (1+o(1)) \frac{t_0^{r-1} p^r}{r!} t_0 = (1+o(1)) \frac{t_0}{n} \stackrel{np=\omega(1)}{=} o(t_0 p) \stackrel{(11)}{=} o(1).\end{aligned}\tag{12}$$

Applying Lemma 7 with $\lambda = \omega_0/2$, we have that with probability at least

$$1 - \exp\left(-\frac{\omega_0^2}{8((1+o(1))n\hat{\pi}(t_c) + \omega_0/6)}\right) \geq 1 - \exp\left(-\frac{\omega_0^2}{10t_0}\right)$$

$M(t_c, n-a) < \omega_0/2$. This together with (10) implies

$$|A(t_c)| \leq t_c - (1+o(1))\omega_0 + \omega_0/2 < t_c.$$

Therefore $T < t_c$ and thus $|A_f| = T < t_c$.

5 Proof of Theorem 2

Before proving Theorem 2 we begin with an observation on $A(t_0)$, the set of infected vertices after the first t_0 steps.

Lemma 8 *Let ω_0 be any function satisfying the conditions $\omega_0 = \omega(\sqrt{a_c})$ and $\omega_0 \leq t_0 - a_c$. If $|A(0)| = a_c + \omega_0$, then with probability at least*

$$1 - \exp\left(-\frac{\omega_0^2}{9.5t_0}\right)$$

we have $T > t_0$ and $|A(t_0)| \geq t_0 + (1+o(1))a_c + \omega_0/2$.

Proof: By the definition of a_c , for every $t \leq t_0$ we have

$$a_c \geq \frac{t - n\hat{\pi}(t)}{1 - \hat{\pi}(t)}.\tag{13}$$

Assume that $M(t, i) > -\omega_0/2$ for every $t \leq t_0$ and $1 \leq i \leq n-a$. First we will show by induction that if $M(t, i)$ satisfies this lower bound, then $T > t_0$. Clearly $T > 0$. Now assume that for some $t \leq t_0 - 1$ we have that $T > t - 1$. Therefore $\pi(t) = \hat{\pi}(t)$. For $t \leq t_0$ we have $\hat{\pi}(t) \leq \hat{\pi}(t_0) \stackrel{(12)}{=} o(1)$ and thus

$$\begin{aligned}|A(t)| &\stackrel{(2)}{=} a + (1 - \pi(t))M(t, n-a) + (n-a)\pi(t) \\ &\stackrel{M(t,i) > -\omega_0/2}{\geq} (1 - \hat{\pi}(t))(a_c + \omega_0) + n\hat{\pi}(t) - (1+o(1))\omega_0/2 \\ &\stackrel{(13)}{\geq} t + (1+o(1))\omega_0/2.\end{aligned}$$

Therefore $|A(t)| > t$ which together with $T > t - 1$ implies $T > t$.

Also note that

$$\begin{aligned} |A(t_0)| &\stackrel{(2)}{=} a + M(t_0, n - a)(1 - \pi(t_0)) + (n - a)\hat{\pi}(t_0) \\ &\stackrel{(12)}{\geq} (1 + o(1))a_c + (1 + o(1))t_0 + (1 + o(1))\omega_0/2. \end{aligned} \quad (14)$$

Let $t_1 := ((r - 1)!/np^r)^{1/(r-1)}$. Then $t_1 \leq t_0$ and so

$$a_c \geq \frac{t_1 - n\hat{\pi}(t_1)}{1 - \hat{\pi}(t_1)}. \quad (15)$$

Also

$$\begin{aligned} \hat{\pi}(t_1) &= \mathbb{P}[\text{Bin}[t_1, p] \geq r] = \sum_{j=r}^{t_0} \binom{t_0}{j} p^j (1 - p)^{t_1 - j} \stackrel{(11)}{=} (1 + o(1)) \frac{t_1^r p^r}{r!} \\ &= (1 + o(1)) \frac{t_1^{r-1} p^r}{r!} t_1 = (1 + o(1)) \frac{t_1}{rn}. \end{aligned} \quad (16)$$

From this and (15)

$$a_c \geq \frac{t_1 - (1 + o(1))t_1/r}{1 - \hat{\pi}(t_1)} = (1 + o(1)) \left(1 - \frac{1}{r}\right) t_1 = \Omega(t_0), \quad (17)$$

which together with (14) and $\omega_0 \leq t_0$ implies

$$|A(t_0)| = t_0 + (1 + o(1))a_c + \omega_0/2.$$

Lemma 7 with $\lambda = \omega_0/2$ implies the result. \square

We will need to establish the size of the giant component in the set of vertices which have at least $r - 1$ neighbours in $Z(t_0)$. For this we first need to determine the number of vertices which have at least $r - 1$ neighbours in $Z(t_0)$.

Lemma 9 *Let $A \subset [n]$ with $|A| = o(n)$. Conditional on $T \geq t_0$, $A(t_0)$ and $Z(t_0)$, with probability $1 - \exp(-\Omega(p^{-1}))$ we have that the number of vertices in $[n] \setminus (Z(t_0) \cup A)$ with at least $r - 1$ neighbours in $Z(t_0)$ is at least $3rp^{-1}/4$.*

Proof: Let X_v be the indicator random variable that a vertex $v \in [n] \setminus (Z(t_0) \cup A)$ has at least $r - 1$ neighbours in $Z(t_0)$ and set $X := \sum_{v \in [n] \setminus (Z(t_0) \cup A)} X_v$. Clearly

$$\mathbb{P}[X_v = 1 | v \in A(t_0)] = 1.$$

Note that if $X_v = 1$ and $v \notin A(t_0)$, then v has exactly $r - 1$ neighbours in $Z(t_0)$ and thus

$$\begin{aligned} \mathbb{P}[X_v = 1 | v \notin A(t_0)] &\geq \mathbb{P}[X_v = 1, v \notin A(t_0)] = \binom{t_0}{r-1} p^{r-1} (1 - p)^{t_0 - r + 1} \\ &\stackrel{(11)}{=} (1 + o(1)) \frac{t_0^{r-1}}{(r-1)!} p^{r-1} = (1 + o(1)) \frac{r}{np}. \end{aligned}$$

Since the set of random variables $\{X_v | v \in [n] \setminus (Z(t_0) \cup A)\}$ are mutually independent, X stochastically dominates the binomial random variable $\hat{X} = \text{Bin}(n - t_0 - |A|, (1 + o(1))r/np)$. Because $t_0 = o(n)$ and $|A| = o(n)$, we have

$$\mathbb{E}[\hat{X}] = (1 + o(1))rp^{-1}$$

and Theorem 3 implies

$$\mathbb{P}[X - \mathbb{E}(\hat{X}) \leq -(1 + o(1))rp^{-1}/4] \leq \exp\left(-(1 + o(1))\frac{r^2p^{-2}}{32rp^{-1}}\right) \leq \exp\left(-\frac{p^{-1}}{32}\right).$$

□

In the following two lemmas we look at the number of vertices with at least r neighbours in a set of order p^{-1} and set of order n . Estimating the probability that a vertex has at least r neighbours in such sets differs significantly and are discussed separately.

Lemma 10 *Let $U, W \subset [n]$ in $G(n, p)$ satisfy $|U| = p^{-1}/2$ and $|W| = o(n)$. With probability at least $1 - \exp(-\Omega(n))$ the number of vertices in $[n] \setminus (U \cup W)$ with at least r neighbours in U is at least $n/(2^r r! \sqrt{e})$.*

Proof: Let Y_v be the indicator random variable that a vertex $v \in [n] \setminus (U \cup W)$ has at least r neighbours in U and set $Y = \sum_{v \in [n] \setminus (U \cup W)} Y_v$. We have that

$$\begin{aligned} \mathbb{P}[Y_v = 1] &= 1 - \sum_{j=0}^{r-1} \binom{p^{-1}/2}{j} p^j (1-p)^{p^{-1}/2-j} \\ &= 1 - (1-p)^{p^{-1}/2} \sum_{j=0}^{r-1} \binom{p^{-1}/2}{j} \left(\frac{p}{1-p}\right)^j \\ &= 1 - (1+o(1))e^{-1/2} \sum_{j=0}^{r-1} \frac{(2p)^{-j}}{j!} p^j \\ &= 1 - (1+o(1))e^{-1/2} \sum_{j=0}^{r-1} \frac{1}{j! 2^j}. \end{aligned}$$

Clearly $1 \geq \mathbb{P}[Y_v = 1] > 1/(2^r r! \sqrt{e})$. Set $\eta := \mathbb{P}[Y_v = 1] - 1/(2^r r! \sqrt{e})$ and note that $\eta = \Omega(1)$. Furthermore the set of random variables $\{Y_v | v \in [n] \setminus (U \cup W)\}$ are mutually independent and $|[n] \setminus (U \cup W)| = (1 + o(1))n$. Therefore, by Theorem 3 we have

$$\mathbb{P}[Y < n/(2^r r! \sqrt{e})] \leq \exp\left(-(1 + o(1))\frac{\eta^2 n^2}{2n}\right) = \exp(-\Omega(n)).$$

□

Lemma 11 *Let $U, W \subset [n]$ in $G(n, p)$ satisfy $|U| = n/(2^r r! \sqrt{e})$. Then with probability $\exp(-\Omega(p^{-1}))$ all but at most p^{-1} vertices in $[n] \setminus (U \cup W)$ have at least r neighbours in U .*

Proof: Let B_v be the indicator random variable that a vertex $v \in [n] \setminus (U \cup W)$ has less than r neighbours in U and set $B := \sum_{v \in [n] \setminus (U \cup W)} B_v$. Since $np = \omega(1)$, we have

$$\mathbb{P}[B_v = 1] = \sum_{j=0}^{r-1} \binom{|U|}{j} p^j (1-p)^{|U|-j} \leq \exp(-|U|p) \sum_{j=0}^{r-1} \frac{(np)^j}{j!} \leq \exp(-|U|p)(np)^{r-1}.$$

Since $|[n] \setminus (U \cup W)| \leq n$ and $np = \omega(1)$, we have

$$\mathbb{E}[B] \leq \exp(-|U|p)n(np)^{r-1} = \exp(-\Omega(np))n^r p^{r-1} = o(p^{-1}).$$

Note that the set of random variables $\{B_v | v \in [n] \setminus (U \cup W)\}$ are mutually independent and therefore, by Theorem 3 we have

$$\mathbb{P}[B > p^{-1}] \leq \exp(-\Omega(p^{-1})).$$

□

Proof of Theorem 2: According to Lemma 8 with probability at least $1 - \exp(-\omega_0^2/(9.5t_0))$ we have that $|A(t)| > t$ for every $t \leq t_0$ and $A(t_0) \geq t_0 + (1 + o(1))a_c + \omega_0/2$. Therefore the process runs for at least t_0 steps and there exists a set $A \subseteq |A(t_0) \setminus Z(t_0)|$ of size $a_c/2 + \omega_0/2$.

Lemma 9 implies that conditional on $A(t_0)$ and $Z(t_0)$ with probability at least $1 - \exp(-\Omega(p^{-1}))$ there is a set of vertices in $[n] \setminus (Z(t_0) \cup A)$ with size at least $3rp^{-1}/4$ where every vertex in the set has at least $r-1$ neighbours in $Z(t_0)$ and select a subset W of these vertices of size exactly $3rp^{-1}/4$. Note that until this point every event depends only on edges with one end in $Z(t_0)$.

According to Theorem 5, with probability $1 - \exp(-\Omega(p^{-1}))$ there is a set $U \subset W$ such that the vertices in U form a connected component and $|U| \geq (1 - \varepsilon)\rho p^{-1}$ for arbitrary $\varepsilon > 0$ independent of n , where ρ is the unique solution of $1 - \rho = \exp(-3\rho r/4)$. Since

$$1 - \rho > \sum_{k=0}^4 \frac{(-3\rho r/4)^k}{k!} > \exp(-3\rho r/4)$$

when $0 < \rho \leq 1/2$, we have $\rho > 1/2$ and thus we have that $|U| \geq p^{-1}/2$. Also this event depends only on the edges with both endpoints in U and thus it is independent of the previous events.

Note that if a vertex in A is connected to a vertex in U , then every vertex in U will become infected. The probability that no vertex in A is connected to any vertex in U is

$$(1-p)^{|U|(a_c+\omega)/2} \leq \exp(-(a_c+\omega)/4).$$

This event depends on edges between A and U and thus it is independent of the previous events.

Now take $U' \subset U$ such that $|U'| = p^{-1}/2$ and denote the set of vertices in $[n] \setminus (W \cup A \cup Z(t_0))$ which have at least r neighbours in U' with B . Since $|W \cup A \cup Z(t_0)| = o(n)$ by Lemma 10 with probability $1 - \exp(-\Omega(n))$ we have that $|B| \geq n/(2^r r! \sqrt{e})$. Note that $B \subset |A_f|$. This event depends only on edges between U' and $[n] \setminus (U \cup W \cup A \cup Z(t_0))$ and thus it is independent of the previous events.

Finally let $B' \subset B$ with $|B'| = n/(2^r r! \sqrt{e})$ and consider the set of vertices in $[n] \setminus (B \cup U \cup Z(t_0) \cup A)$ which contain at least r neighbours in B' . Note that all of these vertices will be infected. By Lemma 11 we have with probability $1 - \exp(-\Omega(p^{-1}))$ that all but at most p^{-1} vertices in $[n] \setminus (B \cup U \cup Z(t_0) \cup A)$

will become infected. Similarly as before this event depends only on edges which we haven't considered previously and thus it is independent of the previous events. Recall that $B \cup U \cup Z(t_0) \cup A \subset A_f$ and thus $|A_f| = (1 + o(1))n$.

Since $p^{-1} = \omega(t_0)$ and $n = \omega(p^{-1})$, we have that the probability that almost every vertex becomes infected is at least

$$1 - \exp\left(-\frac{\omega_0^2}{10t_0}\right) - \exp\left(-\frac{a_c + \omega_0}{4}\right).$$

□

References

- [AF14] H. Amini and N. Fountoulakis. Bootstrap percolation in power-law random graphs. *Journal of Statistical Physics*, 155(1):72–92, 2014.
- [Ami10] H. Amini. Bootstrap percolation in living neural networks. *J. Stat. Phys.*, 141(3):459–475, 2010.
- [BBDCM12] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 364(5):2667–2701, 2012.
- [BR] B. Bollobás and O. Riordan. Exploring hypergraphs with martingales. arXiv:1403.6558.
- [CL02] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, 6(2):125–145, 2002.
- [CL06] F. Chung and L. Lu. Concentration inequalities and martingale inequalities: a survey. *Internet Math.*, 3(1):79–127, 2006.
- [CLR79] J. Chalupa, P. Leath, and G. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C: Solid State Physics*, 12:L31–L35, 1979.
- [ELP⁺] H. Einarsson, J. Lengler, K. Panagiotou, F. Mousset, and A. Steger. Bootstrap percolation with inhibition. arXiv:1410.3291.
- [ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.
- [FSS02] L. Fontes, R. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Comm. Math. Phys.*, 228:495–518, 2002.
- [JŁTV12] S. Janson, T. Łuczak, T. Turova, and T. Vallier. Bootstrap percolation on the random graph $G_{n,p}$. *Ann. Appl. Probab.*, 22(5):1989–2047, 2012.
- [ST85] G.-P. Scalia-Tomba. Asymptotic final-size distribution for some chain-binomial processes. *Adv. in Appl. Probab.*, 17(3):477–495, 1985.